

University of Groningen

## Visual metaphors on anaesthesia monitors do not improve anaesthetists' performance in the operating theatre

van Amsterdam, K.; Cnossen, F.; Ballast, A.; Struys, M. M. R. F.

*Published in:*  
British Journal of Anaesthesia

*DOI:*  
[10.1093/bja/aes516](https://doi.org/10.1093/bja/aes516)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2013

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

van Amsterdam, K., Cnossen, F., Ballast, A., & Struys, M. M. R. F. (2013). Visual metaphors on anaesthesia monitors do not improve anaesthetists' performance in the operating theatre. *British Journal of Anaesthesia*, 110(5), 816-822. <https://doi.org/10.1093/bja/aes516>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Visual metaphors on anaesthesia monitors do not improve anaesthetists' performance in the operating theatre

K. van Amsterdam<sup>1\*</sup>, F. Cnossen<sup>2</sup>, A. Ballast<sup>1</sup> and M. M. R. F. Struys<sup>1</sup>

<sup>1</sup> Department of Anesthesiology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, Postbus 30 001, 9700 RB Groningen, The Netherlands

<sup>2</sup> Department of Artificial Intelligence, University of Groningen, Nijenborgh 9, 9747 AG Groningen, The Netherlands

\* Corresponding author. E-mail: k.van.amsterdam@umcg.nl

## Editor's key points

- A monitor interface that displays data metaphorically with normal values and trends was evaluated.
- Data interpretation from a standard monitor and the new one  $\pm$  trends was compared by 32 clinicians.
- No significant differences were found in time to diagnosis and accuracy between the monitors.
- Visual metaphors did not improve performance and 40% of the complications were identified incorrectly.

**Background.** Previous research using a metaphorical anaesthesia monitor, where dimensions of rectangles proportionally represent 30 patient variable values, showed improved performance in diagnosing adverse events compared with the standard monitor. Steady-state values were represented by a frame around each rectangle. We developed a similar metaphorical anaesthesia interface, but instead of presenting four relatively simple complications, we presented 10 complications of various levels of difficulty. Our simplified monitor presented variables that anaesthetists and trainees suggested as being essential for diagnosis.

**Methods.** Thirty-two anaesthetists and anaesthesia trainees participated in the monitoring task. Three types of monitors were presented: standard monitor, metaphorical monitor, and metaphorical monitor with trend arrows emphasizing the direction of change. The subjects were presented with screenshots of the three monitor types displaying anaesthesia-related complications. They were asked to indicate treatment method and diagnosis for the displayed complication.

**Results.** No significant differences were found in time to diagnosis and accuracy between the metaphorical and standard monitor. There were also no differences between trend and no-trend monitors. Forty per cent of the complications were identified incorrectly.

**Conclusions.** Visual metaphors on anaesthesia monitors do not improve anaesthetists' performance in the operating theatre. Since all complications in this study were identifiable based on monitor values alone, it seems feasible to develop a decision support system (DSS) based on these values. We suggest that a DSS could support the anaesthetist by calling attention to diagnoses that may not be considered.

**Keywords:** decision support systems, clinical; diagnostic errors; monitoring, physiological; pattern recognition, visual

Accepted for publication: 8 November 2012

The proportion of anaesthesia-related incidents due to human error is found to be more than 60%.<sup>1–3</sup> Many of these incidents can be attributed to the physician failing to recognize a pattern in the patient's clinical information.<sup>4</sup> Small deviations in patient variables can evolve into incidents causing harm to the patient.<sup>5,6</sup> However, the presentation of patient variables in current anaesthesia monitors is not optimal for fast detection and interpretation of complications.<sup>7–9</sup> For example, numerical values and curves need to be interpreted one-by-one to obtain a diagnosis.<sup>10</sup> An integrated representation of patient variables in a graphical display has been shown to increase the accuracy and speed of diagnosis.<sup>6,11</sup> The integrated view supports the non-analytical, pattern-recognizing nature of diagnosing, especially when the display highlights the physiological

relationships between the variables.<sup>12</sup> A promising example of such a display uses rectangles to proportionally represent 30 different patient variable values.<sup>6</sup> The steady-state value for each variable was represented by a frame around each rectangle. Because heights and widths of the rectangles changed proportionally with values of patient variables, deviations from normal values are easily detectable. In a patient simulation, anaesthetists were presented with four critical events: blood loss, inadequate paralysis with spontaneous ventilation, cuff leak, and depletion of soda lime. Their results showed a significant effect in detection time and identification time of these events.

However, this monitor displays variables not routinely measured, such as fluid balance.<sup>6</sup> Inspired by this study,<sup>6</sup> we therefore developed a similar metaphorical anaesthesia

interface (MAI), but adapted the design to only show variables that are measured during anaesthesia. We used structured interviews with anaesthetists and anaesthesia trainees to identify preferred variables to include in the design. For example, we included information about the direction in which values change, by adding trend arrows, to emphasize these changes, because it was suggested that would be useful.

We aimed to compare the diagnostic and treatment performance of a standard monitor and our new monitor. In the previous study,<sup>6</sup> four relatively simple complications were presented but we presented subjects with 10 complications of differing complexity. We hypothesized that the subjects would recognize the complications faster with the metaphorical monitor compared with the standard monitor. In addition, we hypothesized that they would recognize the complications faster in monitors with trend arrows compared with those without trend information. Furthermore, we hypothesized increased recognition performance in combined standard and metaphorical monitors compared with single monitors.

## Methods

### Metaphorical display design

In consultation with five anaesthetists, five anaesthesia trainees, and five nurse anaesthetists of the Department of Anaesthesiology of the University Hospital Groningen, we developed a graphical display based on the interface described by Michels and colleagues.<sup>6</sup> Based on the perceived need for accentuating the direction of change in the measured values, we developed two versions of the display: the 'normal' MAI (Fig. 1), and a monitor with trend arrows showing the direction of change (tMAI) (Fig. 2).

The anaesthetists indicated that nine patient variables should be presented, which we grouped according to the physiological system they represented: the respiratory and the vascular system, specified at the top of the screen by

two icons. Inspiratory O<sub>2</sub> (steady state: 45%), expiratory CO<sub>2</sub> (4.3 kPa), respiratory rate (15 resp. min<sup>-1</sup>), tidal volume (450 ml), PEEP (2 cm H<sub>2</sub>O), and PAW (18 cm H<sub>2</sub>O) pressures were shown for the respiratory system, and SpO<sub>2</sub> (100%), heart rate (75 beats min<sup>-1</sup>), and arterial pressure (125/83 mm Hg) for the vascular system. Each variable was presented as a rectangle, in a black frame that represented the steady-state value for this particular patient. Variables below this value did not completely fill up the frame, while variables above it spilled over the top of the frame; variables at the steady-state value fitted exactly. Inspiratory O<sub>2</sub> and expiratory CO<sub>2</sub> were presented in a stacked manner, because that was the preference of the majority of interviewees. Heart rate and arterial pressure were stacked, as a physiological relationship was suggested. A change of height in one of the stacked variables (e.g. arterial pressure) resulted in a higher fragment of that variable in the stack. This means an increase in the height of that particular variable and an equal increase in the height of the total stack (i.e. the other variable in the stack is 'pushed' upwards). The steady state and the changed values for the variables were determined by one of the authors, an experienced anaesthetist with 33 yr experience.

### Trend

In our group of professionals, there was no consensus about the value of trend information in monitoring. We decided to add trend arrows in two monitor conditions. These trend arrows indicated the speed and direction of change for each cardiovascular variable typical for the simulated complication.

### Experimental methods

#### Subjects

Members of the Department of Anaesthesiology at the UMCG Groningen participated in the experiment: 16 anaesthetists (mean age=44.3 yr; sd=9.9 yr; mean experience: 15.4 yr;

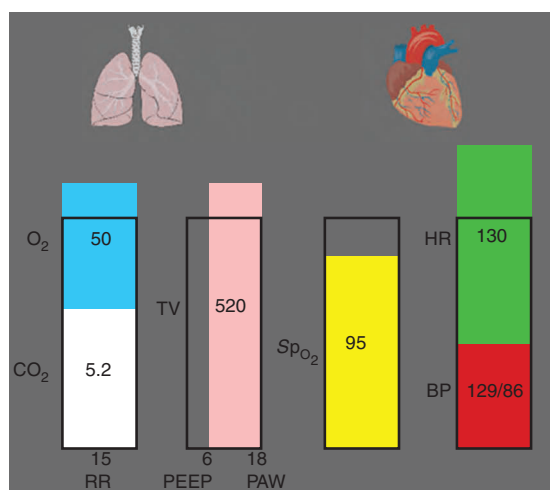


Fig 1 Metaphorical anaesthesia interface (MAI).

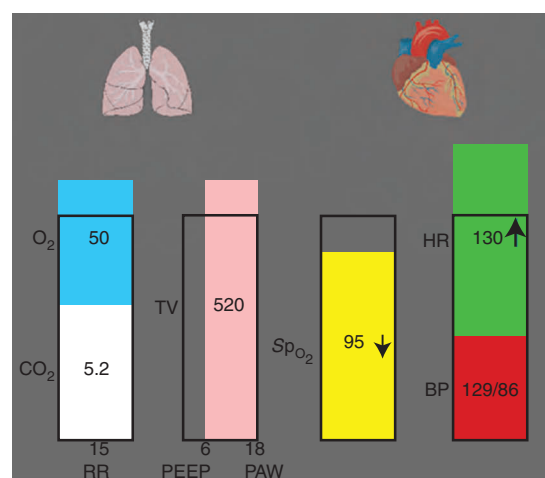


Fig 2 Metaphorical interface with trend information (tMAI).

$SD=9.9$  yr) and 16 anaesthesia trainees (mean age=32.7 yr;  $SD=2.6$  yr; mean experience=3.7 yr;  $SD=2.2$  yr). Clinicians who had been consulted on the design of the graphical display were excluded from participation in the experiment.

### Stimuli

Screenshots of the monitors were presented to the subjects. Five different monitor types and combinations were used: standard, metaphorical, metaphorical+standard, metaphorical with trends, and metaphorical with trends+standard. The screenshot presented a specific anaesthesia-related complication in a male patient (weight 70 kg, height 1.78 m). The complications and corresponding treatments used are shown in Table 1.

### Experimental design

The time to diagnosis was calculated as the time between the presentation of the complication and the subject pressing the 'I know' button. After pressing the button, the screenshot immediately disappeared from the screen to prevent subjects from re-checking the monitor while answering the questions. On the following screen, subjects had to choose the correct diagnosis from a list of possible complications. After that, subjects selected the treatment of their choice from a list of possible treatments.

A within-subject design was used: all subjects were presented with all five monitor types, presented within five blocks. The order of the blocks (monitors) was determined using a Latin-square design. Each experimental block

started with five practices, followed by 11 experimentals (10 complications and one distractor). The distractor trial was added to avoid subjects predicting the complication of the last trial. The complications in each set were presented in a predefined random order. The order of complications in each set was fixed for all subjects.

In the practices, subjects were acquainted with the monitors and the task. They were instructed to select a diagnosis and the appropriate treatment for each complication as quickly and accurately as possible. The experiment lasted 45–60 min. Afterwards, no feedback was given to participants about their performance of the task.

### Apparatus

The experiment was run on an Apple MacBook Pro, Intel Core 2 Duo 2.26 GHz, 2 GB DDR3 SDRAM, 13.3 in glossy TFT LED backlit display (1280×800 pixels) at 80% brightness. The input device for the experiment was an external two-button mouse. The experimental setting was created in Java J2SE 5 (licensed by Sun Microsystems).

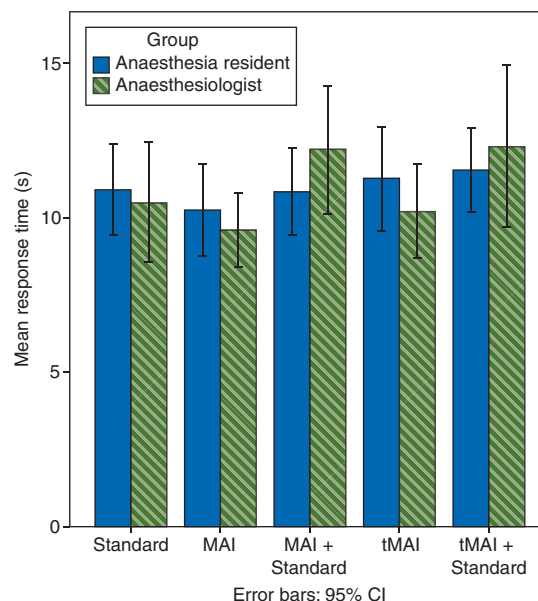
### Data collection and analysis

During the experiment, all subjects' responses were logged. The number of correct identifications was taken as a measure of diagnostic performance. Time to and accuracy of diagnosis and treatment were analysed using repeated-measures analysis of variance (ANOVA) (using SPSS 16.0 Software; SPSS Inc., Chicago, IL, USA).

Power analysis was conducted using G\*Power Software (Heinrich-Heine University, Düsseldorf, Germany). To obtain an effect size of  $f=0.27$  and a power of 96.1% at a 0.05 significance level, we included 32 participants in our design.

**Table 1** The complications presented in the experiment, with their associated appropriate treatments. \*Although temperature could be an indicator for tachycardia with sepsis, we did not include this in the interface. In our preceding pilot study, subjects did not select temperature as an invaluable variable for diagnosing complications

Complication	Appropriate treatment
Air embolism	100% oxygen or vasopressors/ cardiotonics
Anaphylaxis	Vasopressors/cardiotonics
Insufficient depth of anaesthesia	Deepen anaesthesia
Bradycardia	Atropine
Diffusion error	Increase $F_{I_{O_2}}$ or increase PEEP or furosemide
Faulty oxygen supply	Check/correct oxygen supply
Hypoventilation	Increase respiratory minute volume, increase inspiration pressure
Tachycardia with sepsis*	Vasopressors/cardiotonics or volume therapy
Tension pneumothorax	Thoracic drain
Ventilation failure	Manual ventilation or increase respiratory minute volume or increase inspiration pressure
Pulmonary secretions (distractor)	



**Fig 3** Response times between groups.

## Results

### Group effects

One-way repeated-measures ANOVA (Huyn–Feldt corrections) showed no significant main effect of group on time to diagnosis (Fig. 3) ( $F_{1,30}=1.03$ ,  $P=0.32$ ), nor on diagnosis accuracy [ $\chi^2(1)=0.21$ ,  $P=0.65$ ]. As the groups did not differ significantly, we collapsed data across both groups in our further analyses.

### Diagnostic accuracy

Overall, participants correctly diagnosed 60% of all complications. A repeated-measures ANOVA carried out on the number of correct identifications of complications for each monitor showed no significant difference in diagnostic accuracy between monitors,  $F_{4,36}=1.37$ ,  $P=0.26$ . Identification scores for the complications varied highly: subjects correctly identified anaphylaxis in only 10% of the cases, while insufficient depth of anaesthesia was always identified correctly. The scores for anaphylaxis, diffusion, air embolism, and tension pneumothorax were <50%.

### Diagnostic times

Diagnostic times on five types of monitors were comparable: standard (mean=10.69 s,  $SD=4.23$  s), MAI (mean=9.90 s,  $SD=4.34$  s), MAI+standard (mean=11.51,  $SD=5.58$  s), tMAI (mean=10.73 s,  $SD=5.22$  s), and tMAI+standard (mean=11.91 s,  $SD=5.41$  s). A one-way repeated-measures ANOVA (Huyn–Feldt corrections) showed no main effect of monitor type on diagnostic time for trials where a subject selected a correct diagnosis,  $F_{4,124}=1.52$ ,  $P=0.20$  (Fig. 4). There was also no main effect of monitor type on diagnostic times when we included trials with incorrect identifications,  $F_{3,6,112}=1.28$ ,  $P=0.28$ .

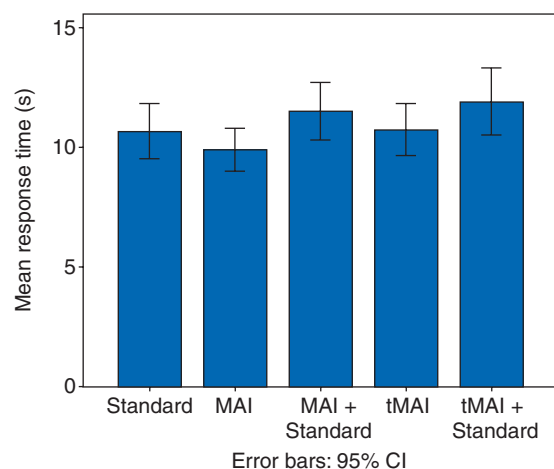


Fig 4 Response times between monitor types.

### Treatment accuracy

In total, there were 1600 treatment decisions; of which, 717 (45%) were inconsistent with the complication presented. Interestingly, almost 30% of all incorrect treatments (203 cases) involved supplying 100% oxygen. In these cases, other treatments were more appropriate.

### Post hoc analyses

Because of the large number of incorrect diagnoses, we performed *post hoc* analyses on the nature of the diagnostic errors. First, we looked at the number of variables in a complication deviating from steady state, which we thought might be indicative of the difficulty of diagnosis. We found a significant relationship between the number of variables deviating from normal and the number of correctly diagnosed complications,  $r=-0.55$ ,  $P$  (two-tailed)=0.042: the higher the number of deviating variables, the lower the diagnostic accuracy.

Because frequency and recency have been found to be important factors in diagnosing,<sup>10</sup> we also compared the complications used in our experiment with those from the Australian Incidence monitoring study,<sup>11</sup> in which 2000 anaesthesia-related incidents were analysed (Table 2). We found that the frequency of anaesthesia-related complications in clinical practice<sup>13</sup> had a relationship to our diagnostic scores [ $r=0.56$ ,  $P$  (two-tailed)=0.037] in that complications with a higher incidence in clinical practice were identified correctly more often in our experiment than less common complications.

## Discussion

The main finding in our controlled study was that diagnostic time and accuracy with our new metaphorical monitor was

**Table 2** Complications: number of deviating variables, incidence rate of 1256 reported incidents,<sup>13</sup> and identification scores in the current experiment

Complication	# Deviating variables	Incidence rate <sup>13</sup>	Identification score in our experiment (%)
Tension pneumothorax	5	6 (0.5%)	33
Ventilation failure	4	155 (12%)	59
Air embolism	4	14 (1.1%)	34
Failure in oxygen supply	3	104 (8.3%)	86
Hypoventilation	3	47 (3.7%)	83
Tachycardia with sepsis	3	31 (2.5%)	74
Anaphylaxis	3	10 (0.8%)	16
Diffusion error	3	7 (0.6%)	24
Bradycardia	2	68 (5.4%)	95
Insufficient anaesthesia	2	No data	99



similar to that with the standard monitor. However, based on the positive results in a previous study,<sup>6</sup> we had hypothesized that anaesthetists and trainees would diagnose complications quicker and more accurately with our metaphorical monitor compared with the standard monitor. That we did not find the same positive effects on diagnostic performance was unexpected.

One explanation of this discrepancy may be the difficulty of the complications presented. In the previous study, four relatively easy complications were presented. We used 10 complications of varying levels of difficulty, some of which proved difficult to diagnose. We also used a static rather than a dynamic monitor, thus depriving the subject of dynamic information. We had included trend arrows in our monitor to indicate dynamic developments. This did not improve diagnostic performance compared with no trend information, so the absence of dynamic information was perhaps not an important factor in the differing results of the two studies. It was nonetheless surprising that trend information did not improve diagnostic performance. Anaesthetists in the initial interviews had indicated that trend information would be a valuable addition to any anaesthesia monitor. Studies on attention to visual cues<sup>10–12</sup> show that trend information increases the saliency of the deviating variables, drawing attention to these variables, and it seems likely that this would improve diagnostic performance.

A further reason for the lack of positive effects on diagnostic performance may lie in the design of our new monitor. The new design is based on the IGAD monitor<sup>6</sup> and on the suggestions from the structured interviews. However, the use of preferences of the interviewees does not guarantee the best usable monitor for all anaesthetists.

Although we cannot definitively explain why our metaphorical monitor did not improve diagnostic performance, the most important finding is possibly the large number of incorrect diagnoses: 40% of the complications were diagnosed incorrectly, irrespective of monitor type. This may be due to the static nature of the monitoring task and the artificial nature of the laboratory setup, which deprived the subjects of contextual information. We did however take care to include only complications that could be diagnosed on the basis of the values on the monitor alone, so it was possible to arrive at the correct diagnosis. The complications were textbook cases with which all anaesthetists should have been familiar. An important advantage of our setup was that there was no real time pressure: although the subjects were asked to respond as fast as possible, without making errors (which is a standard instruction in this type of research), the task itself was self-paced and they could take as long as needed to arrive at a diagnosis. Although this may be different from clinical practice and is thus less realistic, it does mean that the incorrect diagnoses were not the result of lack of thinking time.

Other studies have also found that diagnostic performance of anaesthetists can be suboptimal. For example, in a

dynamic simulation study,<sup>12</sup> anaphylaxis was not recognized by 60% of the subjects, being confused with tachycardia (17% of diagnoses given) and pneumothorax (17%). In our study, anaphylaxis was also identified poorly (only 16%), and was also often confused with tachycardia (37%) and pneumothorax (20%). This suggests that our subjects' performance was not unusual but in accordance with what others have found.

The *post hoc* analysis showed that there was a pattern in the type of errors committed. First, when the values of one complication overlapped with those of another complication, they were easily confused (e.g. confusion between air embolism, tachycardia, and anaphylaxis, which differ only in CO<sub>2</sub> values, was common, Table 3). Secondly, there seems to be an effect of the number of variables deviating from normal and the *amount* of deviation from normal values on diagnosing performance. The more variables deviating from normal, the more difficult it was to diagnose it correctly (Table 2). We speculate that this relates to the limited capacity of working memory, making it difficult to hold and compare many variables in working memory. Lastly, the frequency of complications in clinical practice affects diagnosis selection in our experiment: complications with a high incidence in clinical practice are more often identified correctly in our experiment than those that are less common. This conforms to findings in research on clinical reasoning, where complications or illnesses that occur more frequent and/or more recent are diagnosed faster than lower frequency or less recent complications or illnesses.<sup>13–14</sup> In effect, when the subjects in our study confuse diagnoses, they select the more frequent diagnosis. For example, in the case of anaphylaxis, which resembles both air embolism and tachycardia, the more frequent complication (tachycardia) is preferred over the less frequent complication (air embolism).

It is assumed that generation of hypotheses (diagnoses) is an automatic process, where more recent or more frequent hypotheses are preferred. After hypothesis generation, a deliberate process starts, in which the generated hypotheses are evaluated against the available evidence.

Our results confirm the effect of clinical practice incidence on diagnosis generation but also suggest that the subjects may have failed to evaluate the generated hypotheses more thoroughly. They had time to (re)consider their initial diagnosis (the more frequent complication) and look for information that would have differentiated it from the actual complication. This bias is not uncommon: for example, in a human patient simulator study,<sup>14</sup> fixation errors or failure to revise a plan in the presence of inconsistent cues were made by 63% of subjects. Because this bias is strong and frequent, to counteract this bias, we suggest that a decision support system (DSS) could be useful. The DSS could use the measured values of the physiological variables and call attention to diagnoses that anaesthetists may overlook: complications that are less frequent; with a large number of deviating variables; and complications that are easily confused. A DSS could have an additional role as it could

**Table 3** Number of incorrect diagnoses

# Correct diagnosis	# Selected diagnoses										Total
	Anaphylaxis	Diffusion error	Air embolism	Tension pneumothorax	Ventilation failure	Tachycardia with sepsis	Hypoventilation	Failure in oxygen supply	Bradycardia	Insufficient anaesthesia	
Anaphylaxis	—	8	4	27	2	49	37	1	0	6	134
Diffusion error	0	—	4	4	3	0	100	9	1	1	122
Air embolism	18	12	—	30	0	34	10	1	0	1	106
Tension pneumothorax	17	6	28	—	1	35	17	2	0	1	107
Ventilation failure	1	1	12	10	—	0	20	21	0	0	65
Tachycardia with sepsis	20	5	0	1	2	—	9	0	0	4	41
Hypoventilation	0	17	0	1	5	0	—	0	0	2	25
Failure in oxygen supply	1	9	2	1	3	0	4	—	0	0	20
Bradycardia	2	1	1	0	0	1	0	0	—	3	8
Insufficient anaesthesia	0	0	0	0	0	0	0	0	0	—	0
Total	59	59	51	74	16	119	197	34	1	18	628

suggest diagnoses before physiological changes develop into an adverse event.

Our study may suggest that numerical monitor data are sufficient for clinical diagnosis when the data are clearly presented or incorporated in a DSS. However, the clinical context may not correspond to the monitor data, because of artifacts or when the picture is not typical, leading to confusion and misinterpretation. Therefore, the presentation of monitor data, in whatever format, should support the anaesthetist's clinical judgement, but never replace it.

In this study, we examined whether a metaphorical monitor improves diagnosing performance of anaesthetists. Even though our subjects were not familiar with the metaphorical monitor, their diagnosing performance was similar to that with the standard monitor. Interestingly, a high number (40%) of complications were identified incorrectly, irrespective of monitor type. Diagnostic errors were to a certain extent predictable. However, they included complications that are less frequent, shared similar variable values with other diagnoses, and had a large number of variables values deviating from normal. Anaesthetists normally use direct observational information in addition to monitor values when making a diagnosis. Although all complications in this study were identifiable on the basis of the monitor values alone, it seems feasible to develop a diagnosing support system based on these values. Such a system could support, although never replace, the anaesthetist's clinical judgement by calling attention to rare, but life-threatening diagnoses that may not have been considered. How this should be interlaced within anaesthetic practice warrants further research.

## Declaration of interest

A.B. is an anaesthesiologist in UMCG Groningen. M.M.R.F.S. is an editor and editorial board member of *BJA*, but was not involved in the handling of this manuscript.

## Funding

This study was funded by the Department of Anesthesiology of UMCG Groningen and the Department of Artificial Intelligence of the University of Groningen.

## References

- Cooper JB, Newbower RS, Kitz RJ. An analysis of major errors and equipment failures in anaesthesia management: considerations for prevention and detection. *Anesthesiology* 1984; **60**: 34–42
- DeAnda A, Gaba DM. Unplanned incidents during comprehensive anaesthesia simulation. *Anesth Analg* 1990; **71**: 77–82
- Chopra V, Bovill JG, Spierdijk J, Koornneef F. Reported significant observations during anaesthesia: a prospective analysis over an 18-month period. *Br J Anaesth* 1992; **68**: 13–17
- Cooper JB, Newbower RS, Long CD, McPeck B. Preventable anaesthesia mishaps: a study of human factors. *Anesthesiology* 1978; **49**: 399–406
- Michels P, Gravenstein D, Westenskow DR. An integrated graphic data display improves detection and identification of critical events during anaesthesia. *J Clin Monit* 1997; **13**: 249–59

- 6 Gaba DM, Maxwell M, DeAnda A. Anesthetic mishaps: breaking the chain of accident evolution. *Anesthesiology* 1987; **66**: 670–6
- 7 Cole WG, Stewart JG. Metaphor graphics to support integrated decision making with respiratory data. *Int J Clin Monit Comput* 1993; **10**: 91–100
- 8 Gurushanthaiah K, Weinger MB, Englund CE. Visual display format affects the ability of anesthesiologists to detect acute physiologic changes. A laboratory study employing a clinical display simulator. *Anesthesiology* 1995; **83**: 1184–93
- 9 Charabati S, Bracco D, Mathieu PA, Hemmerling TM. Comparison of four different display designs of a novel anaesthetic monitoring system, the 'integrated monitor of anaesthesia (IMA™)'. *Br J Anaesth* 2009; **103**: 670–7
- 10 Tappan JM, Daniels J, Slavin B, Lim J, Brant R, Ansermino JM. Visual cueing with context relevant information for reducing change blindness. *J Clin Monit Comput* 2009; **23**: 223–32
- 11 Webb RK, Currie M, Morgan CA, et al. The Australian Incident Monitoring Study: an analysis of 2000 incident reports. *Anaesth Intensive Care* 1993; **21**: 520–8
- 12 Schwid HA, O'Donnell D. Anesthesiologists' management of simulated critical incidents. *Anesthesiology* 1992; **76**: 495–501
- 13 Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974; **185**: 1124–31
- 14 Weber E, Böckenholt U, Hilton D, Wallace B. Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. *J Exp Psychol Learn Mem Cogn* 1993; **19**: 1151–64

Handling editor: C. S. Reilly